Package ‘rccdates’
August 29, 2016

Type Package
Title Date Functions for Swedish Cancer Data
Version 1.0.0
Date 2016-07-01
Author Erik Bulow
Maintainer Erik Bulow <erik.bulow@rccvast.se>
Description Identify, convert and handle dates as used within the Swedish cancer register and associated cancer quality registers in Sweden. Especially the cancer register sometimes use nonstandard date variables where day and/or month can be "00" or were the date format is a mixture of "%Y-%m-%d", "%Y%m%d" and "%y%V" (two digit year and week number according to ISO 8601, which is not completely supported by R). These dates must be approximated to valid dates before being used in for example survival analysis. The package also includes some convenient functions for calculating "lead times" (relying on ‘difftime’) and introduce a "year" class with relevant S3-methods to handle yearly cohort.
License GPL-2
Depends R (>= 2.10)
Imports rccmisc
Suggests testthat, knitr, rmarkdown
BugReports https://bitbucket.com/cancercentrum/rccdates/issues
SysDataCompression xz
RoxygenNote 5.0.1
VignetteBuilder knitr
NeedsCompilation no
Repository CRAN
Date/Publication 2016-07-01 16:33:32

R topics documented:

as.Dates ................................................................. 2
as.year ................................................................. 3
as.Dates

Converting all potential dates in a data.frame

Description

This function takes a data.frame, checks all its variables for potential dates and transform potential
date variables to dates.

Usage

as.Dates(x, ...)

## S3 method for class 'data.frame'
as.Dates(x, progress_bar = TRUE, ...)

## Default S3 method:
as.Dates(x, rtr = TRUE, yyww = rtr, yyyyww = rtr,
         yyyy0000 = rtr, yyyymm00 = rtr, ask = FALSE, name = "x", ...)

Arguments

- **x**
  - a data.frame, numeric vector or character vector, possibly with potential date
    variable(s)

- **...**
  - arguments passed to exceed_threshold

- **progress_bar**
  - Should a progress bar be printed to show the progress of the call? TRUE by
default.

- **rtr**
  - Should dates in formats found in RTR be recognised as dates. This controls
    yyww, yyyyww, yyyy0000, yyyymm00 but each of them can also be set indi-
    vidually. TRUE by default.

- **yyww, yyyyww**
  - Should entries of the old yyww/yyyyww-formats be recognised as dates? Same
    as rtr by default.

- **yyyy0000, yyyymm00**
  - Should entries with unknown month and/or day (set to zero) be accepted as
dates? If yes, dates in the form "yyyy0000" will be set to "yyyy0701" and
"yyyymm00" to "yyyymm15". Same as rtr by default.

- **ask, name**
  - arguments passed to exceed_threshold
Details

The function only recognise dates in the standard format used in INCA and Rockan, hence ’ week number sometimes used for older data). We hope that this older format can be omitted in the future since it adds some unclarity if the vector is really a date or not). Dates in other formats (for example as parts of social security numbers, as part of a comment, or as a time stamp) will not be treated as dates.

Value

the given input with potantial date variable(s) converted to date(s) and other variable(s) intact

Slow call

Note that the function call can be very slow for big data sets (each individual cell is checked). You probably only run this function once per dataset and could hopefully live with this drawback :-)

See Also

as.Date

Examples

# Let's say we have a data.frame with one date column and one column that is almost 
# (but not exactly) a date:

test_data <- data.frame(
    not_date = c("19121212", "1912-12-12", "2014-01-01", "121212"),
    date = c(19121212, "1912-12-12", "2014-01-01", "6405")
)

as.Dates(test_data) # Only recognizes the "date" column as date

# Inform that the "not_date" column might also be a date candidate if fixed:
as.Dates(test_data, threshold = .5)

# Force the not_date column to date

as.Dates(test_data, threshold = .5, force = TRUE)

---

as.year

Year vectors

Description

Creates or coerces objects of type year.
Usage

as.year(x)

## S3 method for class 'Date'
as.year(x)

## Default S3 method:
as.year(x)

is.year(x)

## S3 method for class 'year'
a - b

width.year(x)

Arguments

x object to be coerced or tested
a, b years to be added or subtracted

Details

The year class also inherits methods from the character class but not from numeric. (It does not make any sense to, for example, multiply two years with each other). There are however methods for subtraction and it is also possible to add an integer to a year etcetera.

Value

Vector of class year and "AsIs" (see function I).

Examples

as.year("2012")
as.year(Sys.Date())
rccmisc::width(c(2012, 2014))

Description

Coerce yyyyw-variable to a valid date.
asDate

Usage

as.yyyww(x, format_check = TRUE)
as.yyyyww(x, format_check = TRUE)

Arguments

x a vector that can be coerced to a date of the form yV (yyww) or YV (yyyyww)
format_check Should the format of x be checked to be in correct format before we try to change it? TRUE by default. Set to FALSE if the format is already checked outside the function to avoid multiple checks (that could be time consuming for big data sets).

Value

a date vector

Description

This is just a wrapper for the standard INCA date format

Usage

asDate(...) 

Arguments

... arguments passed to as.Date

Examples

asDate(c("", "", "2014-10-10"))
lt  

*Lead time from one date to another*

**Description**

Lead time from one date to another

**Usage**

```r
lt(fromL toL neg = naL as = "numeric")
leadtime(from, to, neg = NA, as = "numeric")
```

**Arguments**

- `from, to` start and stop dates (in formats that can be coerced by as.Dates).
- `neg` default value for negative lead times. NULL means that negative lead times are kept as is. NA by default (changes negative values to NA). Any numerical value is accepted
- `as` name of the class to be assigned to x. Default is numeric but could also be set to for example difftime or integer.

**Value**

A vector of class as (numeric by default).

**Examples**

```r
lt(from = Sys.Date(), to = Sys.Date() + 10)
```

---

**rccdates**  

*Date functions for RCC (INCA and Rockan format)*

**Description**

See as.Dates.

**Author(s)**

Erik Bulow
Index

- .year (as.year), 3
  as.Date, 3, 5
  as.Dates, 2, 6
  as.year, 3
  as.yyww, 4
  as.yyyyww (as.yyww), 4
  asDate, 5
  difftime, 6
  exceed_threshold, 2
  I, 4
  integer, 6
  is.year (as.year), 3
  leadtime (lt), 6
  lt, 6
  rccdates, 6
  rccdates-package (rccdates), 6
  width.year (as.year), 3